

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic) nucleic search, using sw model

Run on: November 20, 1999, 21:34:44 : Search time 50.03 Seconds
(without alignments)
9471.593 Million cell updates/sec

Title: US-09-126-945-1
Perfect score: 1894
Sequence: 1 gtctgactctctcccgagcac.....ataaagactactagagactg 1894

Scoring table: OLIGODX_NUC

Searched: 311585 seqs, 125096042 residues

Database: N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.1	2073	1	Q55149
2	20	1.1	2667	1	T47198
3	20	1.1	2064	1	T37087
4	20	1.1	1070	1	X24412
5	19	1.0	2721	1	N50114
6	19	1.0	1503	1	Q53160
7	19	1.0	5931	1	T04821
8	19	1.0	346	1	T21147
9	19	1.0	22481	1	T11658
10	19	1.0	9048	1	T43225
11	19	1.0	818	1	V32010
12	18	1.0	3931	1	N80777
13	18	1.0	773	1	N82363
14	18	1.0	3635	1	O41652
15	18	1.0	600	1	O78962
16	18	1.0	2562	1	O68817
17	18	1.0	4780	1	O94253
18	18	1.0	1472	1	O98497
19	18	1.0	3931	1	T58758
20	18	1.0	3931	1	T80599
21	18	1.0	3931	1	T94752
22	18	1.0	1472	1	T97323
23	18	1.0	3931	1	V01083
24	18	1.0	1211	1	V34266
25	18	1.0	1804	1	V34267
26	18	1.0	836	1	V80634
27	18	1.0	1144	1	X02117
28	18	1.0	2033	1	X02111
29	18	1.0	1144	1	X02107
30	18	1.0	1930	1	X02108
31	17	0.9	1312	1	N91078
32	17	0.9	920	1	N80315
33	17	0.9	813	1	N80314
34	17	0.9	5212	1	N80317
35	17	0.9	2366	1	O10190
36	17	0.9	2274	1	N60465
37	17	0.9	1312	1	N40079
38	17	0.9	1710	1	O24075
39	17	0.9	15894	1	O40480
40	17	0.9	2938	1	O50644
41	17	0.9	1890	1	O50662
42	17	0.9	1860	1	O50632
43	17	0.9	1408	1	O45301

ALIGNMENTS

c	44	17	0.9	3501	1	Q90529	Rat SIII 110 kDa s
c	45	17	0.9	1800	1	Q91769	Coding sequence of
ALIGNMENTS							
RESULT 1							
ID	Q55149	standard; cDNA to mRNA; 2073 BP.					
AC	Q55149;						
DR	11-JUL-1994	(first entry)					
DE	Adenovirus E1A-F gene.						
KW	Adenovirus; cancer; ets oncogene; HeLa cell; enhancer core sequence; methylation; ds.						
OS	Human adenovirus.						
FT	key	Location/Qualifiers					
FT	cds	844..1311					
FT		/*tag= a					
FT		/note= "Claimed sequence"					
PN	J05328975-A.						
PD	14-DEC-1993.						
PR	02-JUN-1992; 165453.						
PR	02-JUN-1992; JP-165453.						
PA	(TAKI) TAKARA SHUZO CO LTD.						
DR	WPI: 94-021923/03.						
DR	P-PSDB: R45451.						
PT	Novel E1A-F gene - for production of adenovirus E1A-F and cancer research						
PS	Claim 1: Page 6; 7pp; Japanese.						
CC	The adenovirus E1A-F gene contains a 473bp open reading frame. The clone comprising the coding sequence was isolated by screening a HeLa cell cDNA library.						
SO	Sequence 2073 BP; 458 A; 635 C; 561 G; 418 T;						
Query Match 1.1%; Score 20; DB 1; Length 2073;							
Best Local Similarity 100.0%; Pred. No. 8.1;							
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
QY	1319	tacgacagctagccgctc	1338				
DB	1030	TACGACAGCTAGCCGCTC	1049				
RESULT 2							
ID	T47198	standard; cDNA; 2667 BP.					
AC	T47198;						
DR	06-APR-1997	(first entry)					
DE	Human ETS2 repressor factor (ERF) cDNA.						
KW	ETS2 repressor factor; ERF; transcriptional repressor; tumour suppressor; tumour; cancer; oncogene; gene therapy; ss.						
OS	Homo sapiens.						
FT	key	Location/Qualifiers					
FT	cds	123..1769					
FT		/*tag= a					
FT		2644..2649					
FT		/*tag= b					
PN	W09639517-A1.						
PD	12-DEC-1996.						
PR	04-JUN-1996; U10177.						
PR	05-JUN-1995; US-469412.						
PA	(USHS) US DEPT HEALTH & HUMAN SERVICES.						
DR	Athanasiaou MA, Beal GJ, Blair DG, Fisher RJ, Mavrothalassitis GJ;						
DR	WPI: 97-043139/04.						
DR	P-PSDB: W07700.						
PT	New DNA encoding ETS2 repressor factor - useful for reducing tumorigenicity, esp. oncogene associated tumour cells						
PS	Claim 3: Page 63-65; 101pp; English.						
CC	A cDNA clone (T47198) codes for human ETS2 repressor factor (ERF) (W07700). It was isolated from a K562 cDNA library using the H1						

CC site of the ETS2 promoter as probe. A related clone (T47199)
 CC coding for an alternatively spliced ERF (W07701) was also isolated.
 CC The ERF gene, which maps to chromosome 19, q1.2-1.3, is the first
 CC member of the ets family to be identified as a transcriptional
 CC repressor in mammalian cells. It can be used to suppress or repress
 CC transcription and to elucidate transcription process and regulation.
 CC The ERF gene also has tumour suppressor activity and can be used to
 CC reduce ets-dependent tumorigenicity associated with v-mos, c-met,
 CC tpv-met, Ha-ras and gag-myb-ets oncogenes. The cDNA can be cloned
 CC for expression of the ERF polypeptide in host cells.
 SQ Sequence 2667 BP; 456 A; 876 C; 825 G; 510 T;

Query Match 1.1%; Score 20; DB 1; Length 2667;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1272 aggtgagccgctgtgggac 1291
 DB 307 AGGTGAGCCGCTGTGGGC 326

RESULT 3
 ID T37087 standard; cDNA to mRNA; 2064 BP.
 DT 25-APR-1997 (first entry)
 DE E1AF matrix metalloproteinase regulator, cDNA.
 KW E1AF; matrix metalloproteinase; regulator; infiltration; cancer;
 KW metastasis; cell; control; antisense; decoy; DNA binding region;
 KW target DNA; ribosome; induction; diagnosis; detection; treatment;
 KW mammary cancer; fibrosarcoma; osteosarcoma; lung cancer; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 1..1389
 FT /*tag= a
 PN W09624379-A1.
 PD 15-AUG-1996; J00016.
 PF 09-JAN-1996; J00016.
 PR 08-FEB-1995; JP-020173.
 PA (TAKI) TAKARA SHUZO CO LTD.
 PI Fujinaga K, Higashino F, Yoshida K;
 DR WPI: 96-384227/38.
 DR P-PSDB: W00167.
 PT Control of cancer cell infiltration by E1AF gene expression
 PT regulation - also diagnosis of cancer by detection of E1AF gene
 PT expression products
 PS Example 5: Pages 38-42; 65pp; Japanese.
 CC The present sequence encodes the E1AF protein, which is a matrix
 CC metalloproteinase regulator. The infiltration and metastasis of
 CC cancer cells can be controlled by regulating the expression and
 CC expression products of the E1AF gene. This may be accomplished by
 CC inducing antisense DNA or RNA for the E1AF gene, a decoy gene
 CC expressing the DNA binding region of the E1AF protein, the target
 CC DNA for the DNA binding region of the E1AF protein or ribosomes
 CC corresponding to the E1AF gene mRNA. Cancer can be diagnosed by
 CC detecting E1AF gene expression products, e.g. E1AF protein or mRNA.
 CC These methods may be used in the treatment and diagnosis of cancer,
 CC e.g. mammary cancer, fibrosarcoma, osteosarcoma, lung cancer, etc.
 SQ Sequence 2064 BP; 427 A; 648 C; 569 G; 420 T;

Query Match 1.1%; Score 20; DB 1; Length 2064;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1319 tacgacaagctgagccgctc 1338
 DB 1108 TACGACAAGCTGAGCCGCTC 1127

RESULT 4
 X24412

ID X24412 standard; cDNA; 1070 BP.
 AC X24412;
 DT 07-JUN-1999 (first entry)
 DE Maize myo-inositol monophosphatase-3 cDNA.
 KW Myo-inositol monophosphatase-3; maize; corn; phytate;
 KW phytic acid; transgenic plant; animal nutrition; feedstuff; food;
 KW ss.
 OS Zea mays.
 FH Key Location/Qualifiers
 FT cds 57..860
 FT /*tag= a

PN W09905298-A1.
 PD 04-FEB-1999; U14657.
 PF 17-JUL-1998; US-085852.
 PR 18-MAY-1998; US-085852.
 PR 22-JUL-1997; US-053371.
 PR 28-JUL-1997; US-053944.
 PR 08-AUG-1997; US-055526.
 PR 11-AUG-1997; US-055446.
 PA (PION-) PIONEER HI-BRED INT INC.
 PI Beach LR, Bowen BA, Martino-Catt SJ, Wang H, Wang X;
 DR WPI: 99-142948/12.
 DR P-PSDB: W97883.
 PT New polynucleotides controlling phytate metabolism in plants -
 PT useful for improving the nutritional content of plants, by enhancing
 PT levels of non-phytate phosphorus, and reducing phytate levels
 PS Claim 1b: Page 81; 86pp; English.
 CC This is the nucleotide sequence of a cDNA clone encoding maize
 CC myo-inositol monophosphatase-3 (see W97883), an enzyme involved
 CC in the metabolism of phytate. The clone was isolated from a maize
 CC immature ear cDNA library. Polynucleotides (see X24400, X24403,
 CC X24407 and X24410-12) encoding maize phosphatidylinositol-3-kinase
 CC (see W97880), myo-inositol 1,3,4-trisphosphate 5/6-kinase (see
 CC W97881), myo-inositol 1-phosphate synthase and myo-inositol
 CC monophosphatase-3, all enzymes involved in phytate metabolism, are
 CC claimed. The invention relates to the use of such genes to reduce
 CC the levels of phytate, and/or increase the levels of non-phytate
 CC phosphorus, in plants used for food or feed. The genes are
 CC especially used to improve the nutritional content of plants such
 CC as corn and soybean. Transgenic plants, and seed produced by them,
 CC are claimed.
 SQ Sequence 1070 BP; 285 A; 240 C; 292 G; 253 T;

Query Match 1.1%; Score 20; DB 1; Length 1070;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 catgtcgagagcagttcc 989
 DB 56 CATGTCGAGAGCAGTTC 75

RESULT 5
 ID N50114 standard; DNA; 2721 BP.
 AC N50114;
 DT 17-OCT-1991 (first entry)
 DE DNA sequence encoding Epstein-Barr virus (EBV) outer surface protein.
 KW Epstein-Barr virus; antigen; vaccine; ss.
 OS Epstein-Barr virus.
 FH Key Location/Qualifiers
 FT mat_peptide 1..2721
 FT /*tag= a
 FT /label= EBV surface protein antigen
 PN EP-151079-A.
 PD 07-AUG-1985; 400141.
 PF 28-JAN-1985; 400141.
 PR 30-JAN-1984; US-575352.
 PR 23-JUL-1984; US-633558.
 PA (UYCH-) UNIV OF CHICAGO.
 PI Kleff E, Tanner J, Hummel M, Belsel C;
 DR WPI: 85-191978/32.

DR P-PSDB: p50073.
 PT New fragment of Epstein-Barr Virus DNA - useful in vector to
 PT express polypeptide for use in prepn. of vaccine against the
 PT virus and for use in diagnosis.
 PS Claim 1: Page 21-23; 26pp; English.
 CC The sequence encodes an outer surface viral protein of EBV, used
 CC to generate antibodies reacting with the surface proteins of
 CC EBV-infected cells, and in the preparation of a vaccine against EBV.
 SO Sequence 2721 BP: 762 A; 876 C; 557 G; 526 T;

Query Match 1.0%; Score 19; DB 1; Length 2721;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 ccagcccccagtgccca 224
 DB 2277 CCAGCCCCCAGTGCCCA 2295

RESULT 6
 053160/C
 ID 053160 standard; cDNA; 1503 BP.
 AC 053160;
 DT 22-JUN-1994 (first entry)
 DE Sequence encoding retinal pigmented epithelium-derived neurotrophic
 DE factor (PEDNF).
 KM Serine protease inhibitor gene family; neurotrophic activity;
 KM tumour therapy; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 117..1373
 FT /*tag- a
 PN W09324529-A.
 PD 09-DEC-1993.
 PF 04-JUN-1993; U05358.
 PR 04-JUN-1992; US-894215.
 PA (UWSC-) UNIV SOUTHERN CALIFORNIA.
 PI Johnson LV, Tombran-link J;
 DR WPI: 93-405734/50.
 DR P-PSDB: R44800.
 PT Purified retinal pigmented epithelium derived neurotrophic factor
 PT - is used for treating tumours, ocular disease or nerve damage or
 PT as serine protease inhibitors for treating e.g. ischaemia, etc.
 PS Claim 24; pages 44-46; 55pp; English.
 CC PEDNF was isolated from cultured retinal pigment epithelium (RPE)
 CC cells. Oligos were constructed from the sequence derived from PEDNF
 CC and used as primers in PCR amplification of a human fetal eye Charon
 CC BS cDNA library to obtain DNA encoding PEDNF. The oligo primers were
 CC constructed against the following peptides: PEDNF 13 - residues 226-
 CC 244 (053161); and PEDNF 2 - residues 107-135 (053162).
 SO Sequence 1503 BP: 373 A; 437 C; 396 G; 297 T;

Query Match 1.0%; Score 19; DB 1; Length 1503;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 gcagccctggcctgggggt 146
 DB 1248 GCAGCCCTGGCCTGGGGGT 1230

RESULT 7
 T04821
 ID T04821 standard; cDNA; 5931 BP.
 AC T04821;
 DT 18-JAN-1996 (first entry)
 DE EBV gp350/220 cDNA.
 KM EBV: gp350; gp220; gp350/gp220; non-splicing variant; vaccine; ds.
 OS Epstein-Barr virus.
 FH Key Location/Qualifiers
 FT cds 1014..3737

FT /*tag- a
 FT signal_peptide 1014..1067
 FT /*tag- b
 FT mat_peptide 1068..3734
 FT /*tag- c
 FT misc_feature 2514..2515
 FT /*tag- d
 FT /*function- splice donor site
 FT /*note- "bases 2513-2517 (AGGT) are replaced by
 FT GTCG in the non-splicing variant"
 FT misc_feature 3105..3106
 FT /*tag- e
 FT /*function- splice acceptor site
 FT /*note- "bases 3104-3107 (AGGT) are replaced by
 FT TGGG in the non-splicing variant"
 FT poly_a_signal 3742..3747
 FT /*tag- f
 FT W09528488-A1.
 PD 26-OCT-1995.
 PF 13-APR-1995; U04611.
 PR 18-APR-1994; US-229291.
 PA (AVIR-) AVIRON.
 PI Jackman WT, Spaete R;
 DR WPI: 95-373802/48.
 DR P-PSDB: R80144.
 PT New DNA encoding a homogeneous gp350 protein - can be used for
 PT preventing and treating Epstein-Barr virus-related diseases or
 PT conditions
 PS Claim 2; Fig. 1; 61pp; English.
 CC The donor and acceptor splice sites of the EBV gene encoding gp350/
 CC 220 are mutated by replacement of native nucleotides by non-native
 CC nucleotides, without altering the encoded amino acid sequence,
 CC resulting in elimination of gp220 prodn. Recombinant homogeneous
 CC gp350, useful in vaccines, is expressed in mammalian or insect cell
 CC hosts.
 SO Sequence 5931 BP: 1453 A; 1782 C; 1437 G; 1259 T;

Query Match 1.0%; Score 19; DB 1; Length 5931;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 ccagcccccagtgccca 224
 DB 3290 CCAGCCCCCAGTGCCCA 3308

RESULT 8
 T22147/C
 ID T22147 standard; cDNA to mRNA; 346 BP.
 AC T22147;
 DT 02-AUG-1996 (first entry)
 DE Human gene signature H0WGS03717.
 KM Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KM human; cloning; mapping; non-biased library; diagnosis; detection;
 KM cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09514772-A1.
 PD 01-JUN-1995.
 PF 11-NOV-1994; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MATS/) MATSUBARA K.
 PI (OKUBA/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1: Page 1055; 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of

CC human genomic DNA. cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 346 BP; 96 A; 100 C; 75 G; 67 T;

Query Match 1.0%; Score 19; DB 1; Length 346;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 gcagccctggctggggt 146
 ||||||||||||||||
 DB 99 GCAGCCTGGCTGGGGT 81

RESULT 9
 T11658/C
 ID T11658 standard; DNA: 22481 BP.
 AC T11658;
 DT 16-JAN-1997 (first entry)
 DE PEDF full length sequence and flanking sequences.
 KW Pigment epithelium-derived factor; PEDF; neuronal cells; neurons;
 KW glial cells; gliastatic; gliosis; central nervous system; CNS;
 KW neurodegenerative disease; injury; neurotrophic; brain injury;
 KW Parkinson's disease; photoreceptor cells; retina; inhibition;
 KW proliferation; immunoassay; antibody; ageing; degenerative disease;
 KW ds.
 OS Homo sapiens.
 PN M09533480-A1.
 PD 14-DEC-1995.
 PF 06-JUN-1995; U07201.
 PR 07-JUN-1994; US-257963.
 PR 30-DEC-1994; US-367841.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Becerra SP, Chader GJ, Schwartz JP, Tanikawa T;
 DR WPI: 96-039966/04.
 DR P-PSDB: R90287.
 PT Use of pigment epithelium derived factor - for enhancing neuronal
 PT cell survival and inhibiting glial cell proliferation, useful, e.g.
 PT in CNS cell culture or to treat neuro-degenerative diseases
 PS Disclosure; Page 100-122; 151pp; English.
 CC Pigment epithelium-derived factor (PEDF) has both neurotrophic and
 CC gliastatic activity, making it useful in cases where neurons die
 CC quickly and glia tend to proliferate (gliosis), e.g. in CNS cell
 CC culture, in neurodegenerative diseases and in CNS injury. The
 CC neurotrophic effect of PEDF is especially useful for enhancing
 CC survival of neuronal cell cultures intended for use in
 CC transplantation. These include cultures of human foetal brain cells
 CC and neural retina and photoreceptor cells. The gliastatic activity
 CC of PEDF can be applied to inhibiting glial cell proliferation in
 CC certain tumours. Antibodies directed against PEDF can be used for
 CC inhibiting PEDF activity or in an immunoassay for determining
 CC levels of PEDF in fluid, cellular or tissue samples e.g for
 CC determining ageing and/or other degenerative diseases.
 SQ Sequence 22481 BP; 5280 A; 5708 C; 6136 G; 5347 T;

Query Match 1.0%; Score 19; DB 1; Length 22481;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 128 gcagccctggctggggt 146
 ||||||||||||||||
 DB 22009 GCAGCCTGGCTGGGGT 21991

RESULT 10
 T43225
 ID T43225 standard; DNA: 9048 BP.
 AC T43225;
 DT 26-FEB-1997 (first entry)
 DE Brassica napus FCA gene.
 KW FCA; flowering; transgenic plant; oilseed rape; ss.
 OS Brassica napus.
 FH Key
 FT cds
 FT Location/Qualifiers
 FT 2468..2470
 FT /*tag= a
 FT /codon-start= 2468..2470
 FT /note= "translation start codon"
 FT
 FT M09638560-A2.
 PD 05-DEC-1996.
 PF 03-JUN-1996; G01332.
 PR 02-JUN-1995; GB-011196.
 PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
 PI Bancroft I, Dean C, Lister CK, MacKnight RC;
 DR WPI: 97-034373/03.
 DR P-PSDB: M06449.
 PT Methods of influencing flowering characteristics of plants - by
 PT administration of FCA protein, DNA or antisense transcripts
 PS Claim 20; Fig 8a; 97pp; English.
 CC The FCA gene (T43225) of Brassica napus codes for a polypeptide
 CC (M06449) able to influence flowering characteristics, partic.
 CC flowering time. It was isolated from a genomic library using
 CC a cDNA clone obtd. from the Arabidopsis FCA gene (T43224). The
 CC gene fully complements a mutation in the Arabidopsis FCA gene
 CC and is thus a fully functional homologue. Timing of flowering of
 CC transgenic plants can be delayed or hastened using FCA sense and
 CC antisense constructs, mutants and alleles. FCA genes can also be
 CC used to isolate FCA homologues from other plant species.
 SQ Sequence 9048 BP; 2643 A; 1643 C; 1713 G; 3049 T;

Query Match 1.0%; Score 19; DB 1; Length 9048;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 772 ggaagcagtgctccatg 790
 ||||||||||||||||
 DB 7595 GCAGCAGTGCTCCATG 7613

RESULT 11
 V32010/C
 ID V32010 standard; cDNA: 818 BP.
 AC V32010;
 DT 23-SEP-1998 (first entry)
 DE Human Rab protein C (HRABC) cDNA.
 KW Human Rab protein C; HRABC; HRAB; HRAB; Rev; HIV-1;
 KW Intracellular vesicular transport; choroideremia; AIDS; cancer;
 KW exocytosis; endocytosis; ss.
 OS Homo sapiens.
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 130..702
 FT /*tag= a
 FT /product= HRABC
 FT
 FT M09818942-A2.
 PD 07-MAY-1998.
 PF 14-OCT-1997; U18581.
 PR 29-OCT-1996; US-741411.
 PA (INCY-) INCYTE PHARM INC.
 PI Au-Young J, Bandman O;
 DR WPI: 98-272232/24.
 DR P-PSDB: W42097.
 PT New isolated human Rab protein(s) - used to develop products for the
 PT diagnosis, prevention and treatment of choroideremia, AIDS and
 PT cancer
 PS Claim 43; Fig 3A-3C; 88pp; English.

CC The present sequence represents the human Rab protein C (HRABC) cDNA
 CC which was first identified in cDNA library clone 1272054 from the
 CC testicular tumour cDNA library TESTTUM02. The invention also claims
 CC for other human Rab protein (HRAB) cDNAs and the HRAB proteins they
 CC encode, namely HRABA (V32008, W42095), HRABD (V32096) and HRABD
 CC (V32011, W42098). The Rab proteins are claimed to be involved in the
 CC regulation of intracellular vesicular transport in both exocytic and
 CC endocytic pathways. As Rab proteins play a role in mediating the
 CC function of a viral gene, Rev, which is essential for replication of
 CC HIV-1 and as they also mediate cell cycle events, the present Rab
 CC proteins are claimed to be useful in the diagnosis, prevention,
 CC or treatment of choroideremia, AIDS and cancer.
 SO Sequence 818 BP; 178 A; 224 C; 273 G; 142 T;

Query Match 1.0%; Score 19; DB 1; Length 818;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 gccagccctggcctgggg 145
 DB 168 GGCAGCCTGGCTGGGG 150

RESULT 12

ID N80777 standard; cDNA; 3931 BP.

AC N80777;

DT 15-OCT-1990 (first entry)

DE cDNA sequence for a murine 4kb clone encoding murine colony stimulating

DE factor-1 (mucsf-1)

KM Murine colony stimulating factor-1; 4kb clone; murine L-929.

OS Mouse

FS Key Location/Qualifiers

FT cds 160..255

FT /tag- a

FT /product-Leader peptide

FT mat_peptide 256..1877

FT /tag- b

FN W08803173-A.

PD 05-MAY-1988.

PF 16-OCT-1987; U02679.

PR 16-APR-1987; US-039657.

PA (CCTU) Celus Corp.

PI Kochs KE, Helsenbeck RF, Kawasaki ES, Ladner MB;

DR WPI: 88-133287/19.

DR P-PSDB: P80360.

PT New forms of colony stimulating factor-1 -

PT used for enhancing effectiveness of immune system and for

PT stimulating prodn. of lymphokine(s)

PS Disclosure: Fig 4-1 to 4-2; 96pp; English.

CC Total mRNA was extracted and purified from murine L-929 cells, and used

CC to construct a cDNA library in lambda gt10. Approximately 1 million phage

CC plaques were probed with a (32) phosphorous single-stranded CSF-17 DNA. A

CC number of phage plaques which hybridised to probes were purified, and two

CC clones, one with a 2kb insert and the other with a 4kb insert, were

CC selected for further study. The nucleotide sequence for the clones are

CC given in n80777 and n80778; 4kb clone begins at nucleotide 24 relative

CC to the human CSF-17 shown in n80775. There is considerable sequence

CC homology with the human "long form" CSF-1-encoding sequence. After the

CC stop codon the nucleotide sequence diverges widely from the human 3'

CC untranslated sequence in pccsf-17 and in the "long form" clones. A

CC recombinant DNA sequence encoding a protein prepn. having CSF-1 activity

CC which is free of DNA encoding other proteins normally found

CC with CSF-1 is claimed.

SO Sequence 3931 BP; 973 A; 1108 C; 1002 G; 848 T;

Query Match 1.0%; Score 18; DB 1; Length 3931;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1740 gccagccctggcctgggg 1757

DB 2938 GACAAGCCACAGCAG 2955

RESULT 13

ID N82363 standard; DNA; 773 BP.

AC N82363;

DT 29-DEC-1990 (first entry)

DE Sequence encoding human granulocyte macrophage colony stimulating factor

DE (GM-CSF)

KM Lymphokine; Interleukin-3; cancer therapy; ss.

OS Homo sapiens.

FS Key Location/Qualifiers

FT cds 9..59

FT /tag- a

FT misc_feature 138..146

FT /tag- b

FT /note-Region R1*

FT misc_feature 168..176

FT /tag- c

FT /note-Region R2*

FT /tag- d

FT mat_peptide 60..443

FN W08805786-A.

PD 11-AUG-1988.

PF 05-FEB-1988; U00335.

PR 06-FEB-1987; US-011794.

PA (GENE-) Genetics Inst.

PI Clark SC, Mong GG, Donahue RE;

DR WPI: 88-235149/33.

DR P-PSDB: P81886.

PT Colony stimulating factors having reduced carbohydrate levels -

PT obtd. by replacing and/or deleting asparagine residues in GM-CSF

PT sequences

PS Disclosure: Table 1, Page 6; 32pp; English.

CC Proteins characterised by possessing GM-CSF-type biological activity

CC and having a specified peptide sequence, except that 1-6 AAs are replaced

CC and/or deleted within regions Asn-27 - Ser-29 and Asn-37 - Thr-39, such

CC that one or both of the regions are completely deleted or replaced by a

CC single AA residue, a dipeptide or a tripeptide sequence other than

CC Asn-X-Ser or Asn-X-Thr, where X is any AA except for Pro is claimed. Also

CC claimed are cDNA encoding proteins. Variants are active CSFs which may

CC be produced in more homogeneous form and which may possess improved

CC pharmacokinetic profiles relative to natural or recombinant GM-CSF.

SO Sequence 773 BP; 210 A; 196 C; 188 G; 179 T;

Query Match 1.0%; Score 18; DB 1; Length 773;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 ccagcagcagcctggg 138

DB 82 CCAGCAGCAGCCTGGG 99

RESULT 14

ID Q41652 standard; DNA; 3635 BP.

AC Q41652;

DT 26-AUG-1993 (first entry)

DE Phosphoglycerate kinase gene including promoter and terminator.

KM PGR; genomic gene; Aspergillus oryzae; PCR; amplification; ss.

OS Aspergillus oryzae.

FS Key Location/Qualifiers

FT promoter 1..1632

FT /tag- a

FT cds 1633..3018

FT /tag- b

FT Intron 1699..1770

FT /tag- c

FT Intron 2192..2251

```

FI      /tag- d
FI      3019.3635
FT      /tag- e
PN      J05095787-A.
PD      20-APR-1993.
PR      04-OCT-1991: 284117.
PR      04-OCT-1991: JP-284117.
PA      (CHOK/) CHOKAN K.
PA      (JOZO-) JOZO SHIGEN KENRYUSHO KK.
DR      WPI: 93-163587/20.
DR      P-PSDB: R36566.
PT      Phospho:glycerate kinase (I) gene promoter - originating from
PT      genomic gene of phospho:glycerate kinase of Aspergillus oryzae
CC      Claim 1; Page 7-9: 24pp: Japanese.
CC      An Aspergillus oryzae genomic library was prep'd. and primers (see
CC      Q41653-55 and Q43379) whose sequences were based on the N-terminal
CC      sequence of purified PKG were used for PCR amplification of the
CC      phosphoglycerate kinase gene, including the promoter, the coding and
CC      the terminator sequences.
SQ      Sequence 3635 BP: 844 A; 915 C; 874 G; 1002 T;

Query Match
Best Local Similarity 1.0%; Score 18; DB 1; Length 3635;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1523 tgctgctgacctcca 1540
DB      1570 TGCTGCTGACCTTCCA 1587

RESULT 15
Q78962/c
ID      Q78962 standard: DNA: 600 BP.
AC      Q78962;
DE      03-AUG-1995 (first entry)
DE      Human immunoglobulin Vh gene #24.
KW      Primer: PCR; amplify; human; immunoglobulin; variable; heavy chain;
KW      cosmid; placenta; vector; pJB81; E.coli; mammalian; ds.
OS      Homo sapiens.
FH      Key
FH      Location/Qualifiers
FT      cds
FT      68..591
FT      /tag- a
FT      /product= human immunoglobulin variable heavy chain
FT      114..198
FT      /tag- b
FT      /cons_splice= 5' site:No, 3' site:Yes
FT      504..506
FT      /tag- c
FT      /note= "miscellaneous signal, does not conform to
FT      terminator or splice site sequence"
PN      W09426895-A.
PD      24-NOV-1994.
PR      10-MAY-1993: J00603.
PR      10-MAY-1993: WO-J00603.
PA      (NIBS ) JAPAN TOBACCO INC.
PI      Honjo T, Matsuda F;
DR      WPI: 95-006791/01.
DR      P-PSDB: R66316.
PT      DNA fragment comprising human immunoglobulin Vh genes - for the
PT      production of human immunoglobulin in mammalian hosts
PS      Claim 33; Page 61-62: 130pp: Japanese.
CC      A series of genes (Q78939-79002) encoding human immunoglobulin variable
CC      heavy chains. The genes were isolated and cloned from a series of cosmid
CC      constructs: Y202; Y103; Y21; Y6; Y24; 3-31; M84; M18 and M131, by PCR
CC      amplification using primers Q78917-38. The genes are subdivided into 5
CC      families of Vh genes. The fragments cover a region of 800 kb. The DNA
CC      fragments were isolated from high molecular weight DNA from human
CC      placenta. The DNA was partially digested with TagI restriction enzyme.
CC      The fragments were separated by gel electrophoresis and 35-45 kb fractions
CC      were collected. The fragments were ligated with ClaI-digested cosmid
CC      vector pJB81. The ligation products were in vitro packed and infected
CC      into E.coli 490A. The fragments were then subcloned by colony

```

```

CC      hybridisation. The Vh genes and the DNA fragments encoding them are
CC      useful in producing human immunoglobulin in mammalian hosts.
SQ      Sequence 600 BP: 154 A; 163 C; 167 G; 116 T;

Query Match
Best Local Similarity 1.0%; Score 18; DB 1; Length 600;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      776 caggtgcagtcacatggtg 793
DB      81 CAGGTGCAGTCACATGTTG 64

Search completed: November 20, 1999, 21:43:54
Job time: 550 sec

```

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 1999, 21:34:44 : Search time 50.03 Seconds

(Without alignments)
9471.593 Million cell updates/sec

Title: US-09-126-945-1

Perfect score: 1894

Sequence: 1 gtctgactctctccacgacac.....ataaagatactagagactg 1894

Scoring table: OLIGODX_NUC

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.1	2073	1	Adenovirus E1A-F 9
2	20	1.1	2667	1	Human ETS2 repress
3	20	1.1	2064	1	Human ETS2 repress
4	20	1.1	1070	1	Maize myo-inositol
5	19	1.0	2721	1	DNA sequence encod
6	19	1.0	1503	1	Sequence encoding
7	19	1.0	5931	1	EBV gp350/220 CDNA
8	19	1.0	346	1	Human gene signalu
9	19	1.0	22481	1	PEPF full length s
10	19	1.0	9048	1	Brassica napus FcA
11	19	1.0	818	1	Human Rab protein
12	18	1.0	3931	1	CDNA sequence for
13	18	1.0	773	1	Neutrophil nitroge
14	18	1.0	3635	1	Sequence encoding
15	18	1.0	600	1	Phosphoglycerate k
16	18	1.0	2562	1	Human immunoglobul
17	18	1.0	4780	1	DNA encoding the A
18	18	1.0	1472	1	Neutrophil nitroge
19	18	1.0	3931	1	Rabbit 50 kDa dyst
20	18	1.0	3931	1	Murine long form C
21	18	1.0	3931	1	Murine 4 kb colony
22	18	1.0	1472	1	Murine colony stim
23	18	1.0	3931	1	Rabbit (50 kDa) dy
24	18	1.0	1211	1	Gene for mouse col
25	18	1.0	1804	1	Human secreted pro
26	18	1.0	836	1	Kidney injury asso
27	18	1.0	1144	1	Human FEN-1 DNA fr
28	18	1.0	2033	1	Human FEN-1 genom
29	18	1.0	1144	1	Human FEN-1 CDNA.
30	18	1.0	1930	1	Mouse FEN-1 CDNA.
31	17	0.9	1312	1	Alpha-1-antitryps
32	17	0.9	920	1	Sequence encoding
33	17	0.9	813	1	Transcription cont
34	17	0.9	5212	1	Cephalosporin anti
35	17	0.9	2366	1	Synthetic sequence
36	17	0.9	2274	1	Sequence encoding
37	17	0.9	1312	1	Sarcosine oxidase.
38	17	0.9	1710	1	Attenuated measles
39	17	0.9	15894	1	Human Hum-F1-1 ge
40	17	0.9	2938	1	Human Hum-F1-1 ge
41	17	0.9	1890	1	Brain factor-2. Is
42	17	0.9	1860	1	CDNA coding a wate
43	17	0.9	1408	1	

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.1	2073	1	Adenovirus E1A-F 9
2	20	1.1	2667	1	Human ETS2 repress
3	20	1.1	2064	1	Human ETS2 repress
4	20	1.1	1070	1	Maize myo-inositol
5	19	1.0	2721	1	DNA sequence encod
6	19	1.0	1503	1	Sequence encoding
7	19	1.0	5931	1	EBV gp350/220 CDNA
8	19	1.0	346	1	Human gene signalu
9	19	1.0	22481	1	PEPF full length s
10	19	1.0	9048	1	Brassica napus FcA
11	19	1.0	818	1	Human Rab protein
12	18	1.0	3931	1	CDNA sequence for
13	18	1.0	773	1	Neutrophil nitroge
14	18	1.0	3635	1	Sequence encoding
15	18	1.0	600	1	Phosphoglycerate k
16	18	1.0	2562	1	Human immunoglobul
17	18	1.0	4780	1	DNA encoding the A
18	18	1.0	1472	1	Neutrophil nitroge
19	18	1.0	3931	1	Rabbit 50 kDa dyst
20	18	1.0	3931	1	Murine long form C
21	18	1.0	3931	1	Murine 4 kb colony
22	18	1.0	1472	1	Murine colony stim
23	18	1.0	3931	1	Rabbit (50 kDa) dy
24	18	1.0	1211	1	Gene for mouse col
25	18	1.0	1804	1	Human secreted pro
26	18	1.0	836	1	Kidney injury asso
27	18	1.0	1144	1	Human FEN-1 DNA fr
28	18	1.0	2033	1	Human FEN-1 genom
29	18	1.0	1144	1	Human FEN-1 CDNA.
30	18	1.0	1930	1	Mouse FEN-1 CDNA.
31	17	0.9	1312	1	Alpha-1-antitryps
32	17	0.9	920	1	Sequence encoding
33	17	0.9	813	1	Transcription cont
34	17	0.9	5212	1	Cephalosporin anti
35	17	0.9	2366	1	Synthetic sequence
36	17	0.9	2274	1	Sequence encoding
37	17	0.9	1312	1	Sarcosine oxidase.
38	17	0.9	1710	1	Attenuated measles
39	17	0.9	15894	1	Human Hum-F1-1 ge
40	17	0.9	2938	1	Human Hum-F1-1 ge
41	17	0.9	1890	1	Brain factor-2. Is
42	17	0.9	1860	1	CDNA coding a wate
43	17	0.9	1408	1	

CC site of the ETS2 promoter as probe. A related clone (747199)
 CC coding for an alternatively spliced ERF (W07701) was also isolated.
 CC The ERF gene, which maps to chromosome 19, q1.2-1.3, is the first
 CC member of the ets family to be identified as a transcriptional
 CC repressor in mammalian cells. It can be used to suppress or repress
 CC transcription and to elucidate transcription process and regulation.
 CC The ERF gene also has tumour suppressor activity and can be used to
 CC reduce ets-dependent tumorigenicity associated with v-mos, c-met,
 CC tpn-met, Ha-ras and gag-myb-ets oncogenes. The cDNA can be cloned
 CC for expression of the ERF polypeptide in host cells.
 SQ Sequence 2667 BP; 456 A; 876 C; 825 G; 510 T;

Query Match 1.1%; Score 20; DB 1; Length 2667;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1272 aggtggccggcgtctggggc 1291
 DB 307 AGGTGGCCGGCTGTGGGC 326

RESULT 3
 T37087
 ID T37087 standard; cDNA to mRNA; 2064 BP.
 AC T37087;
 DE 25-APR-1997 (first entry)
 KW E1AF matrix metalloproteinase regulator, cDNA.
 KW E1AF; matrix metalloproteinase; regulator; infiltration; cancer;
 KW metastasis; cell; control; antisense; decoy; DNA binding region;
 KW target DNA; ribosome; induction; diagnosis; detection; treatment;
 KW mammary cancer; fibrosarcoma; osteosarcoma; lung cancer; ds.
 OS Homo sapiens.
 FH Key
 FT cds 1.1389 Location/Qualifiers
 FT /*tag= a
 PN W09624379-A1.
 PD 15-AUG-1996;
 PF 09-JAN-1996; J00016.
 PR 08-FEB-1995; JP-020173.
 PA (TAKI) TAKARA SHUZO CO LTD.
 PI Fujinaga K, Higashino F, Yoshida K;
 DR WPI: 96-384227/38.
 DR P-PSDB: W00167.
 PT Control of cancer cell infiltration by E1AF gene expression
 PT regulation - also diagnosis of cancer by detection of E1AF gene
 PT expression products
 PS Example 5; Pages 38-42; 65pp; Japanese.
 CC The present sequence encodes the E1AF protein, which is a matrix
 CC metalloproteinase regulator. The infiltration and metastasis of
 CC cancer cells can be controlled by regulating the expression and
 CC expression products of the E1AF gene. This may be accomplished by
 CC inducing antisense DNA or RNA for the E1AF gene, a decoy gene
 CC expressing the DNA binding region of the E1AF protein, the target
 CC DNA for the DNA binding region of the E1AF protein or ribosomes
 CC corresponding to the E1AF gene mRNA. Cancer can be diagnosed by
 CC detecting E1AF gene expression products, e.g. E1AF protein or mRNA.
 CC These methods may be used in the treatment and diagnosis of cancer,
 CC e.g. mammary cancer, fibrosarcoma, osteosarcoma, lung cancer, etc.
 SQ Sequence 2064 BP; 427 A; 648 C; 569 G; 420 T;

Query Match 1.1%; Score 20; DB 1; Length 2064;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1319 taagacaagctgacgcgcctc 1338
 DB 1108 TACGACAAGCTGACCGCTC 1127

RESULT 4
 X24412

ID X24412 standard; cDNA; 1070 BP.
 AC X24412;
 DT 07-JUN-1999 (first entry)
 DE Maize myo-inositol monophosphatase-3 cDNA.
 KW Myo-inositol monophosphatase-3; maize; corn; phytate;
 KW phytic acid; transgenic plant; animal nutrition; feedstuff; food;
 KW ss.
 OS Zea mays.
 FH Key
 FT cds Location/Qualifiers
 FT 57..860
 FT /*tag= a

PN W09905298-A1.
 PD 04-FEB-1999;
 PF 17-JUL-1998; U14657.
 PR 18-MAY-1998; US-085852.
 PR 22-JUL-1997; US-053371.
 PR 28-JUL-1997; US-053944.
 PR 08-AUG-1997; US-055526.
 PR 11-AUG-1997; US-055446.
 PA (PION-) PIONEER HI-BRED INT INC.
 PI Beach LR, Bowen BA, Martino-Catt SJ, Wang H, Wang X;
 DR WPI: 99-142948/12.
 DR P-PSDB: W97883.
 PT New polynucleotides controlling phytate metabolism in plants -
 PT useful for improving the nutritional content of plants, by enhancing
 PT levels of non-phytate phosphorus, and reducing phytate levels
 PS Claim 1b; Page 81; 86pp; English.
 CC This is the nucleotide sequence of a cDNA clone encoding maize
 CC myo-inositol monophosphatase-3 (see W97883), an enzyme involved
 CC in the metabolism of phytate. The clone was isolated from a maize
 CC immature ear cDNA library. Polynucleotides (see X24400, X24403,
 CC X24407 and X24410-12) encoding maize phosphatidylinositol-3-kinase
 CC (see W97880), myo-inositol 1,3,4-trisphosphate 5/6-kinase (see
 CC W97881), myo-inositol 1-phosphate synthase and myo-inositol
 CC monophosphatase-3, all enzymes involved in phytate metabolism, are
 CC claimed. The invention relates to the use of such genes to reduce
 CC the levels of phytate, and/or increase the levels of non-phytate
 CC phosphorus, in plants used for food or feed. The genes are
 CC especially used to improve the nutritional content of plants such
 CC as corn and soybean. Transgenic plants, and seed produced by them,
 CC are claimed.
 SQ Sequence 1070 BP; 285 A; 240 C; 292 G; 253 T;

Query Match 1.1%; Score 20; DB 1; Length 1070;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 970 catgctggaagagcaattcc 989
 DB 56 CATGTCGAGACAGCAGTCC 75

RESULT 5
 N50114
 ID N50114 standard; DNA; 2721 BP.
 AC N50114;
 DT 17-OCT-1991 (first entry)
 DE DNA sequence encoding Epstein-Barr virus (EBV) outer surface protein.
 KW Epstein-Barr virus; antigen; vaccine; ss.
 OS Epstein-Barr virus.
 FH Key
 FT mat_peptide Location/Qualifiers
 FT 1..2721
 FT /*tag= a
 FT /label= EBV surface protein antigen
 PN EP-151079-A.
 PD 07-AUG-1985;
 PF 28-JAN-1985; 400141.
 PR 30-JAN-1984; US-575352.
 PR 23-JUL-1984; US-633558.
 PA (UYCH-) UNIV OF CHICAGO.
 PI Kleif E, Tanner J, Hummel M, Belsel C;
 DR WPI: 85-191978/32.

DR P-PSDB: P50073.
 PT New fragment of Epstein-Barr Virus DNA - useful in vector to
 PT express polypeptide for use in prepn. of vaccine against the
 PT virus and for use in diagnosis.
 PS Claim 1: Page 21-23: 26pp; English.
 CC The sequence encodes an outer surface viral protein of EBV, used
 CC to generate antibodies reacting with the surface proteins of
 CC EBV-infected cells, and in the preparation of a vaccine against EBV.
 SO Sequence 2721 BP: 762 A: 876 C: 557 G: 526 T;

Query Match 1.0%; Score 19; DB 1: Length 2721;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 ccagagcccccagtgagccaa 224
 DB 2277 CCAGGCCCCCACTGCGCCAA 2295

RESULT 6

O53160/c
 ID O53160 standard; cDNA: 1503 BP.
 AC O53160;
 DT 22-JUN-1994 (first entry)
 DE Sequence encoding retinal pigmented epithelium-derived neurotrophic
 DE factor (PEDNF).
 KW Serine protease inhibitor gene family; neurotrophic activity;
 KW tumour therapy; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT cds 117..1373
 FT /tag- a

PN W09324529-A.
 PD 09-DEC-1993.
 PE 04-JUN-1993: U05358.
 PR 04-JUN-1993: US-894215.
 PA (UYSC-) UNTV SOUTHERN CALIFORNIA.
 PI Johnson LV; Tombran-link J;
 DR WPI: 93-405734/50.
 DR P-PSDB: R44800.

PT Purified retinal pigmented epithelium derived neurotrophic factor
 PT - is used for treating tumours, ocular disease or nerve damage or
 PT as serine protease inhibitors for treating e.g. ischaemia, etc.
 PS Claim 24: pages 44-46: 55pp; English.
 CC PEDNF was isolated from cultured retinal pigment epithelium (RPE)
 CC cells. Oligos were constructed from the sequence derived from PEDNF
 CC and used as primers in PCR amplification of a human fetal eye Cheron
 CC BS cDNA library to obtain DNA encoding PEDNF. The oligo primers were
 CC constructed against the following peptides: PEDNF 13 - residues 226-
 CC 244 (O53161); and PEDNF 2 - residues 107-135 (O53162).
 SO Sequence 1503 BP: 373 A: 437 C: 396 G: 297 T;

Query Match 1.0%; Score 19; DB 1: Length 1503;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 gcagccctggagctgggggt 146
 DB 1248 GCAGCCCTGGGCTGGGGGT 1230

RESULT 7

T04821
 ID T04821 standard; cDNA: 5931 BP.
 AC T04821;
 DT 18-JAN-1996 (first entry)
 DE EBV gp350/220 cDNA.
 DE EBV gp350/220 cDNA.
 KW EBV: gp350: gp220: gp350/gp220: non-splicing variant; vaccine: ds.
 OS Epstein-Barr virus.
 FH Key Location/Qualifiers
 FT cds 1014..3737

FT signal_peptide /tag- a 1014..1067
 FT /tag- b 1068..3734
 FT mat_peptide /tag- c 2514..2515
 FT misc_feature /tag- d
 FT /function- splice donor site
 FT /note- "bases 2513-2517 (AAGT) are replaced by
 FT GTCA in the non-splicing variant"
 FT misc_feature /tag- e 3105..3106
 FT /function- splice acceptor site
 FT /note- "bases 3104-3107 (AGGT) are replaced by
 FT TGA in the non-splicing variant"
 FT poly_a_signal /tag- f 3742..3747
 FT /tag- f

PN W09528488-A1.
 PD 26-OCT-1995.
 PE 13-APR-1995: U04611.
 PR 18-APR-1994: US-229291.
 PA (AVIR-) AVIRON.
 PI Jackman WT, Spaete R;
 DR WPI: 95-373802/48.
 DR P-PSDB: R80144.
 PT New DNA encoding a homogeneous gp350 protein - can be used for
 PT preventing and treating Epstein-Barr virus-related diseases or
 PT conditions
 PS Claim 2: Fig.1: 61pp; English.
 CC The donor and acceptor splice sites of the EBV gene encoding gp350/
 CC 220 are mutated by replacement of native nucleotides by non-native
 CC nucleotides, without altering the encoded amino acid sequence,
 CC resulting in elimination of gp220 prodn. Recombinant homogeneous
 CC gp350, useful in vaccines, is expressed in mammalian or insect cell
 CC hosts.
 SO Sequence 5931 BP: 1453 A: 1782 C: 1437 G: 1259 T;

Query Match 1.0%; Score 19; DB 1: Length 5931;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 ccagagcccccagtgagccaa 224
 DB 3290 CCAGGCCCCCACTGCGCCAA 3308

RESULT 8

T22147/c
 ID T22147 standard; cDNA to mRNA: 346 BP.
 AC T22147;
 DT 02-AUG-1996 (first entry)
 DE Human gene signature HUMGS03717.
 KW Gene signature; messenger RNA; mRNA: relative abundance; frequency;
 KW human: cloning; messenger RNA: relative abundance; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN W09514772-A1.
 PD 01-JUN-1995.
 PE 11-NOV-1994: J01916.
 PR 12-NOV-1993: JP-355504.
 PA (MATS/) MATSUBARA K.
 PA (OKUB/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR WPI: 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1: Page 1055: 2245pp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T26837 and which is able to hybridise to part of

CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (esp. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.
 SQ Sequence 346 BP; 96 A; 100 C; 75 G; 67 T;

Query Match 1.0%; Score 19; DB 1; Length 346;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 128 gcagccctggcctggcctg 146
 DB 99 GCAGCCTGGCCTGGGCT 81

RESULT 9
 T11638/c
 ID T11658 standard; DNA; 22481 BP.
 AC T11658:
 DE 16-JAN-1997 (first entry)
 DT PEDF full length sequence and flanking sequences.
 KW Pigment epithelium-derived factor; PEDF; neuronal cells; neurons;
 KW glial cells; gliastatic; gliosis; central nervous system; CNS;
 KW neurodegenerative disease; injury; neurotrophic; brain cells;
 KW Parkinson's disease; photoreceptor cells; retina; inhibition;
 KW proliferation; immunosay; antibody; aging; degenerative disease;
 KW ds.
 OS Homo sapiens.
 PN W09533480-A1.
 PD 14-DEC-1995.
 PF 06-JUN-1995; 007201.
 PR 07-JUN-1994; US-257963.
 PR 30-DEC-1994; US-367841.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Becerra SP, Chader GJ, Schwartz JP, Tanikaki T;
 DR WPI: 96-038966/04.
 P-PSDB: R90287.
 PT Use of pigment epithelium derived factor - for enhancing neuronal
 PT cell survival and inhibiting glial cell proliferation, useful, e.g.
 PT in CNS cell culture or to treat neuro-degenerative diseases
 PS Disclosure: Page 100-122; 151pp; English.
 CC Pigment epithelium-derived factor (PEDF) has both neurotrophic and
 CC gliastatic activity, making it useful in cases where neurons die
 CC quickly and glia tend to proliferate (gliosis), e.g. in CNS cell
 CC culture, in neurodegenerative diseases and in CNS injury. The
 CC neurotrophic effect of PEDF is especially useful for enhancing
 CC survival of neuronal cell cultures intended for use in
 CC transplantation. These include cultures of human foetal brain cells
 CC and neural retina and photoreceptor cells. The gliastatic activity
 CC of PEDF can be applied to inhibiting glial cell proliferation in
 CC certain tumours. Antibodies directed against PEDF can be used for
 CC inhibiting PEDF activity or in an immunoassay for determining
 CC levels of PEDF in fluid, cellular or tissue samples e.g for
 CC determining ageing and/or other degenerative diseases.
 SQ Sequence 22481 BP; 5280 A; 5708 C; 6136 G; 5347 T;

Query Match 1.0%; Score 19; DB 1; Length 22481;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 128 gcagccctggcctggcctg 146
 DB 22009 GCAGCCTGGCCTGGGCT 21991

RESULT 10
 ID T43225 standard; DNA; 9048 BP.
 AC T43225:
 DE 26-FEB-1997 (first entry)
 KW Brassica napus FCA gene.
 KW FCA; flowering; transgenic plant; oilseed rape; ss.
 OS Brassica napus.
 FH Key
 FT cds
 FT Location/Qualifiers
 FT 2468..2470
 FT /*tag= a
 FT /codon_start= 2468..2470
 FT /note= "translation start codon"
 PN W09638560-A2.
 PD 05-DEC-1996.
 PF 03-JUN-1996; G01332.
 PR 02-JUN-1995; GB-011196.
 PA (INNE-) INNES CENT INNOVATIONS LTD JOHN.
 PI Bancroft I, Dean C, Lister CK, MacKnight RC;
 DR WPI: 97-034373/03.
 P-PSDB: W06449.
 PT Methods of influencing flowering characteristics of plants - by
 PT administration of FCA protein, DNA or antisense transcripts
 PS Claim 20; Fig 8a; 97pp; English.
 CC The FCA gene (T43225) of Brassica napus codes for a polypeptide
 CC (W06449) able to influence flowering characteristics, partic.
 CC flowering time. It was isolated from a genomic library using
 CC a cDNA clone obtd. from the Arabidopsis FCA gene (T43224). The
 CC gene fully complements a mutation in the Arabidopsis FCA gene
 CC and is thus a fully functional homologue. Timing of flowering of
 CC transgenic plants can be delayed or hastened using FCA sense and
 CC antisense constructs, mutants and alleles. FCA genes can also be
 CC used to isolate FCA homologues from other plant species.
 SQ Sequence 9048 BP; 2643 A; 1643 C; 1713 G; 3049 T;

Query Match 1.0%; Score 19; DB 1; Length 9048;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 772 ggagcaggtgcagtcacatg 790
 DB 7595 GCAGCAGTGCAGTCACATG 7613

RESULT 11
 V32010/c
 ID V32010 standard; cDNA; 818 BP.
 AC V32010:
 DT 23-SEP-1998 (first entry)
 DE Human Rab protein C (HRABC) cDNA.
 KW Human Rab protein C; HRABC; HRAB; HRAB; Rev; HIV-1;
 KW Intracellular Vesicular Transport; chorioideremia; AIDS; cancer;
 OS Homo sapiens.
 FH Key
 FT Location/Qualifiers
 FT CDS
 FT 130..702
 FT /*tag= a
 FT /product= HRABC
 PN W09618942-A2.
 PD 07-MAY-1998.
 PF 14-OCT-1997; 018581.
 PR 29-OCT-1996; US-741411.
 PA (INCY-) INCYTE PHARM INC.
 PI Au-Young J, Bandman O;
 DR WPI: 98-272323/24.
 P-PSDB: W42097.
 PT New isolated human Rab protein(s) - used to develop products for the
 PT diagnosis, prevention and treatment of chorioideremia, AIDS and
 PT cancer
 PS Claim 43; Fig 3A-3C; 88pp; English.

CC The present sequence represents the human Rab protein C (HRAB) cDNA
 CC which was first identified in cDNA library clone 1272054 from the
 CC testicular tumour cDNA library TESTTUT02. The invention also claims
 CC for other human Rab protein (HRAB) cDNAs and the HRAB proteins they
 CC encode, namely HRAB (V32008, W42095), HRAB (V32009, W42096) and HRAB
 CC (V32011, W42098). The Rab proteins are claimed to be involved in the
 CC regulation of intracellular vesicular transport in both exocytic and
 CC endocytic pathways. As Rab proteins play a role in mediating the
 CC function of a viral gene, Rev, which is essential for replication of
 CC HIV-1 and as they also mediate cell cycle events, the present Rab
 CC proteins are claimed to be useful in the diagnosis, prevention,
 CC or treatment of choroideremia, AIDS and cancer.

SO Sequence 818 BP; 178 A; 224 C; 273 G; 142 T;

Query Match 1.0%; Score 19; DB 1; Length 818;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 127 ggcagccctggcggcg 145

Db 168 GGCAGCCCTGGCGTGGCG 150

RESULT 12

ID N80777 standard; cDNA; 3931 BP.

AC N80777;

DT 15-OCT-1990 (first entry)

DE cDNA sequence for a murine 4kb clone encoding murine colony stimulating

DE factor-1 (mcsf-1)

KM Murine colony stimulating factor-1; 4kb clone; murine L-929.

OS Mouse

FS Key Location/Qualifiers

FT cds 160..255

FT mat_peptide 256..1877

FT

PN W08803173-A.

PD 05-MAY-1988.

PF 16-OCT-1987; U02679.

PR 16-APR-1987; US-039657.

PA (Cetus) Cetus Corp.

PI Kochs KE, Halenbeck RF, Kawasaki ES, Ladner MB;

DR WPI: 88-133247/19.

DR P-PSDB: P80360.

PT New forms of colony stimulating factor-1 -

PT used for enhancing effectiveness of immune system and for

PT stimulating prodn of lymphokine(s)

PS Disclosure: Fig 4-1 to 4-2; 96pp; English.

CC Total mRNA was extracted and purified from murine L-929 cells, and used

CC to construct a cDNA library in lambda gt10. Approximately 1 million phage

CC plaques were probed with a (32) phosphorous single-stranded CSF-17 DNA. A

CC number of phage plaques which hybridised to probes were purified, and two

CC clones, one with a 2kb insert and the other with a 4kb insert, were

CC selected for further study. The nucleotide sequence for the clones are

CC given in n80777 and n80778. 4kb clone begins at nucleotide 24 relative

CC to the human CSF-17 shown in n80775. There is considerable sequence

CC homology with the human "long form" CSF-1-encoding sequence. After the

CC stop codon the nucleotide sequence diverges widely from the human 3'

CC untranslated sequence in pCSF-17 and in the "long form" clones. A

CC recombinant DNA sequence encoding a protein prepn. having CSF-1 activity

CC which is free of DNA encoding other proteins normally found

CC with CSF-1 is claimed.

SO Sequence 3931 BP; 973 A; 1108 C; 1002 G; 848 T;

Oy 1740 gacaaagccacagcgag 1757

Db 2938 GACAAAGCCACAGCAG 2955

RESULT 13

ID N82363 standard; DNA; 773 BP.

AC N82363;

DT 29-DEC-1990 (first entry)

DE Sequence encoding human granulocyte macrophage colony stimulating factor

DE (GM-CSF)

KM Lymphokine; Interleukin-3; cancer therapy; ss.

OS Homo sapiens.

FS Key

FT cds 9..59

FT misc_feature 138..146

FT misc_feature 168..176

FT misc_feature 188..196

FT misc_feature 200..208

FT misc_feature 212..220

FT misc_feature 224..232

FT misc_feature 236..244

FT misc_feature 248..256

FT misc_feature 260..268

FT misc_feature 272..280

FT misc_feature 284..292

FT misc_feature 296..304

FT misc_feature 308..316

FT misc_feature 320..328

FT misc_feature 332..340

FT misc_feature 344..352

FT misc_feature 356..364

FT misc_feature 368..376

FT misc_feature 380..388

FT misc_feature 392..400

FT misc_feature 404..412

FT misc_feature 416..424

FT misc_feature 428..436

FT misc_feature 440..448

FT misc_feature 452..460

FT misc_feature 464..472

FT misc_feature 476..484

FT misc_feature 488..496

FT misc_feature 500..508

FT misc_feature 512..520

FT misc_feature 524..532

FT misc_feature 536..544

FT misc_feature 548..556

FT misc_feature 560..568

FT misc_feature 572..580

FT misc_feature 584..592

FT misc_feature 596..604

FT misc_feature 608..616

FT misc_feature 620..628

FT misc_feature 632..640

FT misc_feature 644..652

FT misc_feature 656..664

FT misc_feature 668..676

FT misc_feature 680..688

FT misc_feature 692..700

FT misc_feature 704..712

FT misc_feature 716..724

FT misc_feature 728..736

FT misc_feature 740..748

FT misc_feature 752..760

FT misc_feature 764..772

FT misc_feature 776..784

Db 2938 GACAAAGCCACAGCAG 2955

RESULT 13

ID N82363 standard; DNA; 773 BP.

AC N82363;

DT 29-DEC-1990 (first entry)

DE Sequence encoding human granulocyte macrophage colony stimulating factor

DE (GM-CSF)

KM Lymphokine; Interleukin-3; cancer therapy; ss.

OS Homo sapiens.

FS Key

FT cds 9..59

FT misc_feature 138..146

FT misc_feature 168..176

FT misc_feature 188..196

FT misc_feature 200..208

FT misc_feature 212..220

FT misc_feature 224..232

FT misc_feature 236..244

FT misc_feature 248..256

FT misc_feature 260..268

FT misc_feature 272..280

FT misc_feature 284..292

FT misc_feature 296..304

FT misc_feature 308..316

FT misc_feature 320..328

FT misc_feature 332..340

FT misc_feature 344..352

FT misc_feature 356..364

FT misc_feature 368..376

FT misc_feature 380..388

FT misc_feature 392..400

FT misc_feature 404..412

FT misc_feature 416..424

FT misc_feature 428..436

FT misc_feature 440..448

FT misc_feature 452..460

FT misc_feature 464..472

FT misc_feature 476..484

FT misc_feature 488..496

FT misc_feature 500..508

FT misc_feature 512..520

FT misc_feature 524..532

FT misc_feature 536..544

FT misc_feature 548..556

FT misc_feature 560..568

FT misc_feature 572..580

FT misc_feature 584..592

FT misc_feature 596..604

FT misc_feature 608..616

FT misc_feature 620..628

FT misc_feature 632..640

FT misc_feature 644..652

FT misc_feature 656..664

FT misc_feature 668..676

FT misc_feature 680..688

FT misc_feature 692..700

FT misc_feature 704..712

FT misc_feature 716..724

FT misc_feature 728..736

FT misc_feature 740..748

FT misc_feature 752..760

FT misc_feature 764..772

FT misc_feature 776..784

Query Match 1.0%; Score 18; DB 1; Length 773;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 121 ccagcagcagcctggg 138

Db 82 CCAGCAGCAGCCTGGG 99

RESULT 14

ID Q41652 standard; DNA; 3635 BP.

AC Q41652;

DT 26-AUG-1993 (first entry)

DE Phosphoglycerate kinase gene including promoter and terminator.

DE PGR: genomic gene; Aspergillus oryzae; PCR; amplification; ss.

OS Aspergillus oryzae.

FS Key

FT promoter 1..1632

FT promoter 1633..3018

FT cds 3019..3018

FT intron 1699..1770

FT intron 2192..2251

```

FT      terminator      /*tag- d
FT      3019. 3635
FT      /*tag- e
PN      J05095787-A.
PD      20-APR-1993.
PF      04-OCT-1991: 284117.
PR      04-OCT-1991: JP-284117.
PA      (CHOK/) CHOKAN K.
PA      (JOZO-) JOZO SHIGEN KENKYUSHO KK.
DR      WPI: 93-163587/20.
DR      P-PSDB: R36566.
PT      Phospho:glycerate kinase (I) gene promoter - originating from
PS      genomic gene of phospho:glycerate kinase of Aspergillus oryzae
PS      Claim 1, Page 7-9: 24pp: Japanese.
CC      An Aspergillus oryzae genomic library was prep'd. and primers (see
CC      A41653-55 and Q42379) whose sequences were based on the N-terminal
CC      sequence of purified PKG were used for PCR amplification of the
CC      phosphoglycerate kinase gene, including the promoter, the coding and
CC      the terminator sequences.
SQ      Sequence 3635 BP; 844 A; 915 C; 874 G; 1002 T;

```

```

Query Match      1.0%; Score 18; DB 1: Length 3635;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1523 tgctgctcctgacctcca 1540
DB      1570 TGCTGCTCTGACCTTCA 1587

```

RESULT 15

Q78962/C
ID Q78962 standard; DNA: 600 BP.

AC Q78962;

DT 03-AUG-1995 (first entry)

DE Human immunoglobulin Vh gene #24.

KW Primer: PCR: amplify: human: immunoglobulin: variable; heavy chain;

OS cosmid; placenta: vector: pJ8B1; E.coli; mammalian; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT cds

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

FT intron

misc-signal
504.506
/*tag- b
/cons-splice- 5' site:No, 3' site:Yes
/*tag- c
/note- "miscellaneous signal, does not conform to
terminator or splice site sequence"

W09426895-A.

24-NOV-1994.

PF 10-MAY-1993: J00603.

PR 10-MAY-1993: WO-J00603.

PA (NISB) JAPAN TOBACCO INC.

PI Honjo T, Matsuda F;

DR WPI: 95-006791/01.

DR P-PSDB: R66316.

PT DNA fragment comprising human immunoglobulin Vh genes - for the

production of human immunoglobulin in mammalian hosts

PS Claim 33; Page 61-62: 130pp: Japanese.

CC A series of genes (Q78939-79902) encoding human immunoglobulin variable

heavy chains. The genes were isolated and cloned from a series of cosmid

constructs: Y202; Y103; Y21; Y61Y24; 3-31; M64; M18 and M31, by PCR

amplification using primers Q78917-38. The genes are subdivided into 5

families of Vh genes. The fragments cover a region of 800 kb. The DNA

fragments were isolated from high molecular weight DNA from human

placenta. The DNA was partially digested with Taqi restriction enzyme.

CC The fragments were separated by gel electrophoresis and 35-45 kb fractions

CC were collected. The fragments were ligated with ClaI-digested cosmid

into E.coli 490A. The fragments were then subcloned by colony

CC hybridisation. The Vh genes and the DNA fragments encoding them are
CC useful in producing human immunoglobulin in mammalian hosts.
SQ Sequence 600 BP; 154 A; 163 C; 167 G; 116 T;

Query Match 1.0%; Score 18; DB 1: Length 600;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 776 caggtgcagtcacatggtg 793
DB 81 CAGGTGCATGCTCATGGTG 64

Search completed: November 20, 1999, 21:43:54
Job time: 550 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 20, 1999, 21:34:44 ; Search time 50.03 seconds
(without alignments)
9471.593 Million cell updates/sec

Title: US-09-126-945-1

Perfect score: 1894

Sequence: 1 gctgactcctccagcac.....ataagatactagagaactg 1894

Scoring table: OLGODX_NUC

Searched: 311585 segs, 125096042 residues

Database : N.Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.1	2073	1	O55149
2	20	1.1	2667	1	T47198
3	20	1.1	2064	1	T37087
4	19	1.1	1070	1	X24412
5	19	1.0	2721	1	N50114
6	19	1.0	1503	1	O53160
7	19	1.0	5931	1	T04821
8	19	1.0	346	1	T22147
9	19	1.0	22481	1	T11658
10	19	1.0	9048	1	T43225
11	19	1.0	818	1	V32010
12	18	1.0	3931	1	N80777
13	18	1.0	773	1	N82363
14	18	1.0	3635	1	O41652
15	18	1.0	600	1	Q78962
16	18	1.0	2562	1	O89817
17	18	1.0	4780	1	O94253
18	18	1.0	1472	1	O98497
19	18	1.0	3931	1	T58758
20	18	1.0	3931	1	T80599
21	18	1.0	3931	1	T94752
22	18	1.0	1472	1	T97323
23	18	1.0	3931	1	V01083
24	18	1.0	1211	1	V34265
25	18	1.0	1804	1	V34267
26	18	1.0	836	1	V80634

```

27 18 1.0 1144 1 X02117 Human FEN-1 DNA fr
28 18 1.0 2033 1 X02111 Human FEN-1 genom
29 18 1.0 1144 1 X02107 Human FEN-1 CDNA.
30 18 1.0 1930 1 X02108 Mouse FEN-1 CDNA.
31 17 0.9 1312 1 N91078 Alpha-1-antitryps
32 17 0.9 920 1 N80315 Sequence encoding
33 17 0.9 813 1 N80314 Sequence encoding
34 17 0.9 5212 1 N80317 Transcription cont
35 17 0.9 2366 1 Q0190 Cephalosporin anti
36 17 0.9 2274 1 N60465 Synthetic sequence
37 17 0.9 1312 1 N40079 Sequence encoding
38 17 0.9 1710 1 Q24075 Sarcosine oxidase.
39 17 0.9 15894 1 Q40480 Attenuated measles
40 17 0.9 2938 1 Q50644 Human Hum-Fli-1 ge
41 17 0.9 1890 1 Q50662 Human Hum-Fli-1 ge
42 17 0.9 1860 1 Q50632 Brain factor-2. Is
43 17 0.9 1408 1 Q45301 RNA coding a wate
44 17 0.9 3501 1 Q90529 Rat S111 110 kDa s
45 17 0.9 1800 1 Q91769 Coding sequence of

```

ALIGNMENTS

RESULT 1

055149 055149 standard: cDNA to mRNA: 2073 BP.

AC 055149: 11-JUL-1994 (first entry)

DE Adenovirus E1A-F gene.

KW Adenovirus; cancer; ets oncogene; HeLa cell; enhancer core sequence;

CC methylation: ds.

OS Human adenovirus.

FT Key Location/Qualifiers

FT cds 844..1311

FT /tag= a /note= "Claimed sequence"

FN J05328975-A.

PD 14-DEC-1993.

PF 02-JUN-1992: 165453.

PR 02-JUN-1992: JP-165453.

PA (TAKI) TAKARA SHUZO CO LTD.

DR WPI: 94-021923/03.

P-PSDB: R45451.

PT Novel E1A-F gene - for production of adenovirus E1A-F and cancer

PS Claim 1: Page 6: 7pp: Japanese.

CC The adenovirus E1A-F gene contains a 473bp open reading frame. The

CC clone comprising the coding sequence was isolated by screening

CC a HeLa cell cDNA library.

SC Sequence 2073 BP: 458 A; 635 C; 561 G; 418 T;

Query Match 1.1%; Score 20; DB 1; Length 2073;

Best Local Similarity 100.0%; Pred. No. 8.1;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1319 tacgacagctgagccgctc 1338

DB 1030 TACGACAGCTGAGCCGCTC 1049

RESULT 2

747198 747198 standard: cDNA: 2667 BP.

AC 747198:

DT 06-APR-1997 (first entry)

DE Human ETS2 repressor factor (ERF) cDNA.

KW ETS2 repressor factor; ERF; transcriptional repressor;

OS tumour suppressor; cancer; oncogene; gene therapy; ss.

FT Key

FT cds Location/Qualifiers

FT 123..1769

FT polyA_signal /*tag= a 2644..2649

FT /*tag= b

FN W09639517-A1.

PD 12-DEC-1996.

PF 04-JUN-1996: U10177.

PR 05-JUN-1995: US-469412.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Athanasios MA, Beal GJ, Blair DG, Fisher RJ, Mavrouthalassitis GJ;

PI Sgouras D N;

DR WPI: 97-043139/04.

P-PSDB: W07700.

PT New DNA encoding ETS2 repressor factor - useful for reducing

PT tumorigenicity ets oncogene associated tumour cells

PS Claim 3, Page 63-65: 10pp: English.

CC A cDNA clone (747198) codes for human ETS2 repressor factor (ERF)

CC (W07700). It was isolated from a K562 cDNA library using the H1

CC site of the ETS2 promoter as probe. A related clone (747199)

CC coding for an alternatively spliced ERF (W07701) was also isolated.

CC The ERF gene, which maps to chromosome 19, q1.2-1.3, is the first

CC member of the ets family to be identified as a transcriptional

CC repressor in mammalian cells. It can be used to suppress or repress

CC transcription and to elucidate transcription process and regulation.

CC The ERF gene also has tumour suppressor activity and can be used to

CC reduce ets-dependent tumorigenicity associated with v-mos, c-met,

CC tpr-met, Ha-ras and gag-myb-ets oncogenes. The cDNA can be cloned

CC for expression of the ERF polypeptide in host cells.

SC Sequence 2667 BP: 456 A; 876 C; 825 G; 510 T;

Query Match 1.1%; Score 20; DB 1; Length 2667;

Best Local Similarity 100.0%; Pred. No. 8;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1272 aggtgagccgctgtgggac 1291

DB 307 AGGTGAGCCCGCTGTGGGAC 326

RESULT 3

T37087 T37087 standard: cDNA to mRNA: 2064 BP.

AC T37087:

DT 25-APR-1997 (first entry)

DE E1A-F matrix metalloproteinase regulator. cDNA.

KW E1A-F; matrix metalloproteinase; regulator; infiltration; cancer;

KW metastasis; cell; control; antisense; decoy; DNA binding region;

KW target DNA; ribosome; induction; diagnosis; detection; treatment;

KW mammary cancer; fibrosarcoma; osteosarcoma; lung cancer; ds.

OS Homo sapiens.

FT Key Location/Qualifiers

FT cds 1..1389

FT /*tag= a

FN W09624379-A1.

PD 15-AUG-1996.

PF 09-JAN-1996: J00016.

PR 08-FEB-1995: JP-020173.

PA (TAKI) TAKARA SHUZO CO LTD.

DR WPI: 96-384227/38.

P-PSDB: W00167.

PT Control of cancer cell infiltration by E1A-F gene expression

PT regulation - also diagnosis of cancer by detection of E1A-F gene

PS Example 5: Pages 38-42: 65pp: Japanese.

CC The present sequence encodes the E1A-F protein, which is a matrix

CC metalloproteinase regulator. The infiltration and metastasis of

CC cancer cells can be controlled by regulating the expression and

CC expression products of the E1A-F gene. This may be accomplished by

CC inducing antisense DNA or RNA for the E1A-F gene, a decoy gene

CC expressing the DNA binding region of the E1A-F protein, the target

CC DNA for the DNA binding region of the E1A-F protein or ribosomes

CC corresponding to the E1A-F gene mRNA. Cancer can be diagnosed by

CC detecting ELAF gene expression products, e.g. ELAF protein or mRNA.
 CC These methods may be used in the treatment and diagnosis of cancer.
 CC e.g. mammary cancer, fibrosarcoma, osteosarcoma, lung cancer, etc.
 SQ Sequence 2064 BP; 427 A; 648 C; 569 G; 420 T;

Query Match 1.1%; Score 20; DB 1; Length 2064;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1319 tagagaaagctagccgc 1338
 |||||
 DB 1108 TAGACAACTGAGCCGCTC 1127

RESULT 4
 X24412 standard; cDNA; 1070 BP.

AC X24412: 07-JUN-1999 (first entry)
 DE Maize myo-inositol monophosphatase-3 cDNA.
 KM Myo-inositol monophosphatase-3; maize; corn; phytate;
 KM phytic acid; transgenic plant; animal nutrition; feedstuff; food;
 OS Zea mays.
 FH Key Location/Qualifiers
 FT CDS 57..860
 ET /*tag= a

MO9905298-A1.
 PD 04-FEB-1999.
 PF 17-JUL-1998; U14657.
 PR 18-MAY-1998; US-083852.
 PR 22-JUL-1997; US-053371.
 PR 28-JUL-1997; US-053944.
 PR 08-AUG-1997; US-055526.
 PR 11-AUG-1997; US-055446.
 PA (PION-) PIONEER HI-BRED INT INC.
 PI Beach LR, Bowen BA, Martino-Calt SJ, Wang H, Wang X;
 DR WPI: 99-142948/12.
 DR P-PSDB: W97883.

PT New polynucleotides controlling phytate metabolism in plants -
 useful for improving the nutritional content of plants, by enhancing
 PT levels of non-phytate phosphorus, and reducing phytate levels
 PS Claim 1b: Page 61, 86pp; English.
 CC This is the nucleotide sequence of a cDNA clone encoding maize
 CC myo-inositol monophosphatase-3 (see W97883), an enzyme involved
 CC in the metabolism of phytate. The clone was isolated from a maize
 CC immature ear cDNA library. Polynucleotides (see X24400, X24403,
 CC X24407 and X24410-12) encoding maize phosphatidylinositol-3-kinase
 CC (see W97880), myo-inositol 1,3,4-triphosphate 5/6-kinase (see
 CC W97881), myo-inositol 1-phosphate synthase and myo-inositol
 CC monophosphatase-3, all enzymes involved in phytate metabolism, are
 CC claimed. The invention relates to the use of such genes to reduce
 CC the levels of phytate, and/or increase the levels of non-phytate
 CC phosphorus, in plants used for food or feed. The genes are
 CC especially used to improve the nutritional content of plants such
 CC as corn and soybean. Transgenic plants, and seed produced by them,
 CC are claimed.
 SQ Sequence 1070 BP; 285 A; 240 C; 292 G; 253 T;

Query Match 1.1%; Score 20; DB 1; Length 1070;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 970 catgctgagagagcattcc 989
 |||||
 DB 56 CATGTCGAGAGCAGATTCC 75

RESULT 5
 NS0114 standard; DNA; 2721 BP.

AC NS0114:
 DT 17-OCT-1991 (first entry)
 DE DNA sequence encoding Epstein-Barr virus (EBV) outer surface protein.
 KM Epstein-Barr virus; antigen; vaccine; ss.
 OS Epstein-Barr virus.

FH Key Location/Qualifiers
 FT mat_peptide 1..2721
 ET /*tag= a
 FT /*label= EBV surface protein antigen

EP-151079-A.
 PD 07-AUG-1985.
 PF 28-JAN-1985; 400141.
 PR 30-JAN-1984; US-575352.
 PR 23-JUN-1984; US-633558.
 PA (UYCH-) UNIT OF CHICAGO.
 PI Kieft E, Tanner J, Hummel M, Belsel C;
 DR WPI: 85-191978/32.
 DR P-PSDB: P50073.

PT New fragment of Epstein-Barr Virus DNA - useful in vector to
 PT express polypeptide for use in prep. of vaccine against the
 PT virus and for use in diagnosis.
 PS Claim 1; Page 21-23; 26pp; English.
 CC The sequence encodes an outer surface viral protein of EBV, used
 CC to generate antibodies reacting with the surface proteins of
 CC EBV-infected cells, and in the preparation of a vaccine against EBV.
 SQ Sequence 2721 BP; 762 A; 876 C; 557 G; 526 T;

Query Match 1.0%; Score 19; DB 1; Length 2721;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 ccagcccccagtgccaa 224
 |||||
 DB 227 CCAGCCCCCAGTGCCAA 2295

RESULT 6
 053160/c
 ID 053160 standard; cDNA; 1503 BP.

DT 22-JUN-1994 (first entry)
 DE Sequence encoding retinal, pigmented epithelium-derived neurotrophic
 DE factor (PEDNF).
 KM Serine protease inhibitor gene family; neurotrophic activity;
 KM tumour therapy; ss.
 OS Homo sapiens.

FH Key Location/Qualifiers
 FT cds 117..1373
 ET /*tag= a

PN WO9324529-A.
 PD 09-DEC-1993.
 PF 04-JUN-1993; U05358.
 PR 04-JUN-1993; US-894215.
 PA (UYSC-) UNIT SOUTHERN CALIFORNIA.
 PI Johnson LV, Tombran-Tink J;
 DR WPI: 93-405734/50.
 DR P-PSDB: R44800.

PT Purified retinal pigmented epithelium derived neurotrophic factor
 PT as serine protease inhibitors for treating e.g. ischemia, etc.
 PS Claim 24; pages 44-46; 55pp; English.
 CC PEDNF was isolated from cultured retinal pigment epithelium (RPE)
 CC cells. Oligos were constructed from the sequence derived from PEDNF
 CC and used as primers in PCR amplification of a human fetal eye Chiron
 CC BS cDNA library to obtain DNA encoding PEDNF. The oligo primers were
 CC constructed against the following peptides: PEDNF 13 - residues 226-
 CC 244 (Q53161); and PEDNF 2 - residues 107-135 (Q53162).
 SQ Sequence 1503 BP; 373 A; 437 C; 396 G; 297 T;

Query Match 1.0%; Score 19; DB 1; Length 1503;
 Best Local Similarity 100.0%; Pred. No. 23;

transplantation. These include cultures of human foetal brain cells and neural retina and photoreceptor cells. The glial cell proliferation of PDF can be applied to inhibiting glial cell proliferation in certain tumours. Antibodies directed against PDF can be used for inhibiting PDF activity or in an immunoassay for determining levels of PDF in fluid, cellular or tissue samples e.g for determining ageing and/or other degenerative diseases.

Sequence 22481 BP; 5280 A; 5708 C; 6136 G; 5347 T;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 22481;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 ggcgcctgagctgggggt 146
Db 22009 GCAGCCCTGGCTGGGGGT 21991

RESULT 10
ID T43225 standard: DNA; 9048 BP.
AC T43225;
DT 26-FEB-1997 (first entry)
DE Brassica napus FCA gene.
KW FCA: flowering; transgenic plant; oilseed rape; ss.
OS Brassica napus.
FH Key location/Qualifiers
FT cds 2468..2470
FT /tag= a
FT /codon_start= 2468..2470
FT /note= "translation start codon"

WO9638560-A2.
PD 05-DEC-1996.
PF 03-JUN-1996; G01332.
PR 02-JUN-1995; GB-011196.
PI (INNE-) INNES CENT INNOVATIONS LTD JOHN.
DR Bancroft I, Dean C, Hester CK, MacKnight RC;
DR WPI: 97-034373/03.
PT Methods of influencing flowering characteristics of plants - by administration of FCA protein; DNA or antisense transcripts
PS Claim 20; Fig 8a; 97pp; English.
CC The FCA gene (T43225) of Brassica napus codes for a polypeptide (W06449) able to influence flowering characteristics, partic. flowering time. It was isolated from a genomic library using a cDNA clone obtd. from the Arabidopsis FCA gene (T4324). The gene fully complements a mutation in the Arabidopsis FCA gene and is thus a fully functional homologue. Timing of flowering of transgenic plants can be delayed or hastened using FCA sense and antisense constructs, mutants and alleles. FCA genes can also be used to isolate FCA homologues from other plant species.
CC Sequence 9048 BP; 2643 A; 1643 C; 1713 G; 3049 T;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 9048;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 772 ggaagcagtgagtcacatg 790
Db 7595 GGAGCAGCTGCAGTCATG 7613

RESULT 11
V32010/c
ID V32010 standard: cDNA; 818 BP.
AC V32010;
DT 23-SEP-1998 (first entry)
DE Human Rab protein C (HRAB).
KW Human Rab protein C; HRAB; HRAB; HRAB; HRAB; Rev. HIV-1; intracellular vesicular transport; choriolidermia; AIDS; cancer; exocytosis; endocytosis; ss.

OS Homo sapiens.
FH Key location/Qualifiers
FT CDS 130..702
FT /tag= a
FT /product= HRAB

WO9818942-A2.
PD 07-MAY-1998.
PF 14-OCT-1997; U18581.
PR 29-OCT-1996; US-741411.
PA (INCY-) INCYTE PHARM INC.
PI Au-Young J, Bandman O;
DR WPI: 98-272232/24.
DR P-PSDB: W42037.

PT New isolated human Rab protein(s) - used to develop products for the diagnosis, prevention and treatment of choriolidermia, AIDS and cancer
PS Claim 43; Fig 3A-3C; 88pp; English.
CC The present sequence represents the human Rab protein C (HRAB) cDNA which was first identified in cDNA incyte clone 1272054 from the testicular tumour cDNA library TESTTUT02. The invention also claims for other human Rab protein (HRAB) cDNAs and the HRAB proteins they encode, namely HRABA (V32008, W42095), HRABB (V32009, W42096) and HRABD (V32011, W42098). The Rab proteins are claimed to be involved in the regulation of intracellular vesicular transport in both exocytic and endocytic pathways. As Rab proteins play a role in mediating the function of a viral gene, Rev, which is essential for replication of HIV-1 and as they also mediate cell cycle events, the present Rab proteins are claimed to be useful in the diagnosis, prevention, or treatment of choriolidermia, AIDS and cancer.
SQ Sequence 818 BP; 178 A; 224 C; 273 G; 142 T;

Query Match
Best Local Similarity 100.0%; Score 19; DB 1; Length 818;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 ggcagccctggctggggg 145
Db 168 GCAGCCCTGGCTGGGGG 150

RESULT 12
ID N80777 standard: cDNA; 3931 BP.
AC N80777;
DT 15-OCT-1990 (first entry)
DE cDNA sequence for a murine 4kb clone encoding murine colony stimulating factor-1 (mucsf-1)

DE factor-1 (mucsf-1)
KW Murine colony stimulating factor-1; 4kb clone; murine L-929.
OS Mouse
FH Key location/Qualifiers
FT cds 160..255
FT /tag= a
FT /product= Leader peptide
FT /tag= b

mat_peptide 256..1877
WO8803173-A.
PD 05-MAY-1988.
PF 16-OCT-1987; U02679.
PR 16-APR-1987; US-039657,
PA (CETU) Cetus Corp.
PI Kohns KE, Halenbeck RF, Kawasaki ES, Ladner MB;
DR WPI: 88-133247/19.
DR P-PSDB: P80360.

PT New forms of colony stimulating factor-1 - used for enhancing effectiveness of immune system and for stimulating prodn. of lymphokine(s)
PS Disclosure: Fig 4-1 to 4-2; 96pp; English.
CC Total mRNA was extracted and purified from murine L-929 cells, and used to construct a cDNA library in lambda gt10. Approximately 1 million phage clones were probed with a (32) phosphorous single-stranded CSF-17 DNA. A number of phage plaques which hybridised to probes were purified, and two clones, one with a 2kb insert and the other with a 4kb insert, were

CC selected for further study. The nucleotide sequence for the clones are given in n80777 and n80778. 4kb clone begins at nucleotide 24 relative to the human CSF-17 shown in n80775. There is considerable sequence homology with the human "long form" CSF-1 encoding sequence. After the stop codon the nucleotide sequence diverges widely from the human 3' unretranslated sequence in pCCSF-17 and in the "long form" clones. A CC recombinant DNA sequence encoding a protein prep. having CSF-1 activity which is free of DNA encoding other proteins normally found with CSF-1 is claimed.

CC Sequence 3931 BP: 973 A; 1108 C; 1002 G; 848 T;

Query Match 1.0%; Score 18; DB 1; Length 3931;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1740 gacaaaggccacagcagc 1757
|||||
DB 2938 GACAAAGCCACAGCAGC 2955

RESULT 13
N82363
ID N82363 standard; DNA: 773 BP.
AC N82363:
DE 29-DEC-1990 (first entry)
DE Sequence encoding human granulocyte macrophage colony stimulating factor (GM-CSF)
KW Lymphokine; Interleukin-3; cancer therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 9..59 /*tag- a
FT misc-feature 138..146 /*tag- b
FT /*note="Region R1" 168..176
FT misc-feature /*tag- c
FT /*note="Region R2" 60..443
FT mat_peptide /*tag- d
FT W08805786-A.
PD 11-AUG-1988.
PF 05-FEB-1988; U00335.
PR 06-FEB-1987; US-011794.
PA (GENE-) Genetics Inst.
PI Clark SC, Wong GG, Donahue RE;
DR WPI: 88-235149/33.
DR P-PSDB; P81886.
PT Colony stimulating factors having reduced carbohydrate levels - obtd. by replacing and/or deleting asparagine residues in GM-CSF sequences
PS Disclosures: Table 1, Page 6; 32pp; English.
CC Proteins characterized by possessing GM-CSF-type biological activity and having a specified peptide sequence, except that 1-6 AAs are replaced and/or deleted within regions Asn-27 - Ser-29 and Asn-37 - Thr-39, such that one or both of the regions are completely deleted or replaced by a single AA residue, a dipeptide or a tripeptide sequence other than Asn-X-Ser or Asn-X-Thr, where X is any AA except for Pro is claimed. Also claimed are cDNA encoding proteins. Variants are active CSFs which may be produced in more homogeneous form and which may possess improved pharmacokinetic profiles relative to natural or recombinant GM-CSF.
SO Sequence 773 BP: 210 A; 196 C; 188 G; 179 T;

Query Match 1.0%; Score 18; DB 1; Length 773;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 ccagcagcagcagcctcgg 138
|||||
DB 82 CCAGCAGCAGCCTCGG 99

RESULT 14
041652
ID 041652 standard; DNA: 3635 BP.
AC 041652:
DE 26-AUG-1993 (first entry)
DE Phosphoglycerate kinase gene including promoter and terminator.
KW PKC; genomic gene; Aspergillus oryzae; PCR; amplification; ss.
OS Aspergillus oryzae.
FH Key Location/Qualifiers
FT promoter 1..1632
FT cds 1633..3018 /*tag- a
FT /*tag- b 1699..1770
FT intron 2192..2251 /*tag- c
FT /*tag- d 3019..3635
FT terminator /*tag- e

PN J05095787-A.
PD 20-APR-1993.
PF 04-OCT-1991; JP-284117.
PR 04-OCT-1991; JP-284117.
PA (CHOK/) CHOKAN K.
DR WPI: 93-163587/20.
DR P-PSDB; R36566.
PT Phosphoglycerate kinase (I) gene promoter - originating from genomic gene of phosphoglycerate kinase of Aspergillus oryzae
PS Claim 1: Page 7-9; 24pp; Japanese.
CC An Aspergillus oryzae genomic library was prepd. and primers (see 041653-55 and 042379) whose sequences were based on the N-terminal sequence of purified PKC were used for PCR amplification of the CC phosphoglycerate kinase gene, including the promoter, the coding and the terminator sequences.
SO Sequence 3635 BP: 844 A; 915 C; 874 G; 1002 T;

Query Match 1.0%; Score 18; DB 1; Length 3635;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1523 tgcgtcctgacctcca 1540
|||||
DB 1570 TGCCTCTGACTTCCA 1587

RESULT 15
078962/c
ID 078962 standard; DNA: 600 BP.
AC 078962:
DE 03-AUG-1995 (first entry)
DE Human immunoglobulin Vh gene #24.
KW Primer: PCR; amplification; human; immunoglobulin; variable; heavy chain; COSmid; placent; vector; pub81; E.coli; mammalian; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 68..591 /*tag- a
FT /*product= human immunoglobulin variable heavy chain 111..198
FT intron /*tag- b
FT /*tag- c 504..506
FT misc-signal /cons-splice= 5' site:No, 3' site:Yes
FT /*tag- c
FT /*note="miscellaneous signal, does not conform to terminator or splice site sequence"

PN W09426895-A.
PD 24-NOV-1994.
PF 10-MAY-1993; J00603.
PR 10-MAY-1993; WO-J00603.
PA (NISB) JAPAN TOBACCO INC.

PI Houjo T, Matsuda F;
 DR WPI: 95-006791/01.
 P-PSDB: R66316.
 PT DNA fragment comprising human immunoglobulin Vh genes - for the
 PT production of human immunoglobulin in mammalian hosts
 PS Claim 33; Page 61-62; 130pp; Japanese.
 CC A series of genes (Q78939-79002) encoding human immunoglobulin variable
 heavy chains. The genes were isolated and cloned from a series of cosmid
 constructs Y202; Y103; Y21; Y6; Y24; 3-31; M84; M18 and M131, by PCR
 amplification using primers Q78917-38. The genes are subdivided into 5
 families of Vh genes. The fragments cover a region of 800 kb. The DNA
 fragments were isolated from high molecular weight DNA from human
 placenta. The DNA was partially digested with TaqI restriction enzyme.
 CC The fragments were separated by gel electrophoresis and 35-45 kb fractions
 were collected. The fragments were ligated with ClaI-digested cosmid
 vector pUB81. The ligation products were in vitro packed and infected
 into E.coli 490A. The fragments were then subcloned by colony
 hybridisation. The Vh genes and the DNA fragments encoding them are
 CC useful in producing human immunoglobulin in mammalian hosts.
 SQ Sequence 600 BP; 154 A; 163 C; 167 G; 116 T;

Query Match 1.08; Score 18; DB 1; Length 600;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 776 caggtcagtcacatggtg 793
 ||||||||||||||||
 DB 81 CAGGTGCACTCCATGCTG 64

Search completed: November 20, 1999, 21:43:54
 Job time: 550 sec